

単眼 RGB カメラを用いた手話の動作解析

Sign language analysis with monocular RGB camera

柳澤 利紀[†]石川 孝明[‡]渡辺 裕^{†‡}Toshiki Yanagisawa[†]Takaaki Ishikawa[‡]Hiroshi Watanabe^{†‡}[†]早稲田大学基幹理工学部情報通信学科[†]School of Fundamental Science and Engineering,
Waseda University[‡]早稲田大学国際情報通信センター[‡]Global Information and Telecommunication Institute,
Waseda University

1. まえがき

手話はろう者にとって重要なコミュニケーション手段であるが、健聴者の多くは手話を知らないため意思の疎通が難しい。この問題を解決するために手話の自動認識システムが望まれてきている。

これまで取り組まれてきた方法には、センサグローブを用いる方法や、Kinect, LeapMotion といった Depth カメラを用いる方法、さらに単眼 RGB カメラと OpenCV を用いる方法などがある。しかし、センサグローブや Depth センサを用いる方法では特別な器具が必要であり、単眼 RGB カメラと OpenCV を用いる方法では肌の色や背景、光量などに左右されやすいといった問題がある。

そこで本研究では、単眼 RGB カメラから高い精度で骨格抽出ができる OpenPose を用いて手話の動作を解析する手法を提案する。

2. OpenPose

OpenPose[1][2]は CVPR2017 にてカーネギーメロン大学 (CMU) の Zhe らによって発表された骨格抽出アルゴリズムである。RGB カメラから多人数の体や顔、手のキーポイントをリアルタイムに抽出できる。

3. LSTM

LSTM は RNN に入力ゲート、忘却ゲート、出力ゲートからなる LSTM ブロックを付けた構造をしており、系列データの分析に利用される。このブロックがメモリのような働きをすることで従来の RNN が苦手としていた長期の依存関係の分析が可能となる。

4. 提案手法

本研究では OpenPose により得られた座標データのうち、表 1 に示す 7 点を処理対象とする。

表 1. 本研究で用いる OpenPose の部位

OpenPose における部位番号	部位
1	首
2	右肩
3	右肘
4	右手首
5	左肩
6	左肘
7s	左手首

首の座標を $P_1 = (P_{1x}, P_{1y})$ 、右肩の座標を $P_2 = (P_{2x}, P_{2y})$ のように表し、首から各部位までのベクトル $\vec{P}_{12}, \vec{P}_{13}, \vec{P}_{14}, \vec{P}_{15}, \vec{P}_{16}, \vec{P}_{17}$ を首から両肩までの長さの平均 $(|\vec{P}_{12}| + |\vec{P}_{15}|)/2$ で割って得られるベクトルの x 座標と y 座標、計 12 点を特徴量とする。

両手手話の動作 5 種類 (お疲れ様、お大事に、危ない、久しぶり、笑う) について、9 人から各動作をおよそ 50 回ずつ集め、上記特徴量を用いて LSTM によって分類を行なった。

5. 実験結果

8 名のデータを学習用に、1 名のデータをテスト用に用いる交差検定を、テスト用に用いる人を変えながら 9 回行った。その結果を表 2 に示す。今回の実験では正解率 95.11% が得られた。

表 2. 交差検定の結果

	動作 1	動作 2	動作 3	動作 4	動作 5
動作 1	445	0	2	4	0
動作 2	2	444	5	0	0
動作 3	53	0	407	0	0
動作 4	3	0	2	420	16
動作 5	16	0	0	7	425

6. むすび

本研究では単眼 RGB カメラを用いた手話の動作解析を OpenPose と LSTM を用いて行い、正解率 95.11% を得た。単語数が増えた場合に手の座標情報も含める必要が出てくるが、特徴量の次元が増大し精度が向上しない問題点がある。

7. 参考文献

[1] Z.Cao, G.Hidalgo, T.Simon, S-E Wei, Y.Sheikh, "Realtime Multi - Person 2D Pose Estimation using Part Affinity Fields," <https://arxiv.org/abs/1812.08008>, 18, Dec, 2018.

[2] Github, inc. CMU-Perceptual-Computing-Lab/openpose, <https://github.com/CMU-Perceptual-Computing-Lab/openpose>, 3, Jan, 2019.